DSTO

AR-010-503

Tests of cognitive, perceptual and sustained attention functions in hot environments

Mark J. Patterson, Nigel A.S. Taylor and Denys Amos

DSTO-TR-0650

19980909 082

APPROVED FOR PUBLIC RELEASE

DEPARTMENT OF DEFENCE

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION

AQF98-12-2255

Tests of Cognitive, Perceptual and Sustained Attention Functions in Hot Environments

Mark J. Patterson, Nigel A.S. Taylor and Denys Amos

Combatant Protection and Nutrition Branch
Aeronautical and Maritime Research Laboratory

DSTO-TR-0650

ABSTRACT

The report reviews the utility of various tests of cognitive function during human performance in hot conditions. The evidence that the thermal environment does impact upon cognitive, perceptual or motor functions is not unequivocal. The lack of consistency in the quantification of ambient conditions, body core and skin temperatures restricts the value of many investigations. Differences in task duration and complexity may lead to disputable conclusions being drawn. Overall, heat stress does appear to impact upon some forms of cognitive and motor performance. Guidelines and procedures for selecting appropriate cognitive, perceptual and sustained attention tests are discussed. Tests suitable for determination of the effects of heat on psychological performance are recommended. Experimental conditions detailing the degree of thermal strain appropriate for cognitive function tests in the heat are described.

RELEASE LIMITATION

Approved for public release

DEPARTMENT OF DEFENCE

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION

Published by

DSTO Aeronautical and Maritime Research Laboratory
PO Box 4331
Melbourne Victoria 3001 Australia

Telephone: (03) 9626 7000
Fax: (03) 9626 7999

AR-010-503
April 1998

Tests of Cognitive, Perceptual and Sustained Attention Functions in Hot Environments

Executive Summary

Conventional wisdom is that heat stress affects the cognitive ability of people operating in hot environments. However, there is some anecdotal evidence to support a view that only some aspects of cognitive performance are modified by thermal strain. This has major implications for ADF operations in northern Australia, especially with the release of new generations of equipment to enhance soldier performance which may impose increased cognitive and physiological loads on the operator.

This report reviews the utility of various tests of cognitive functions during human performance in hot conditions. The evidence that the thermal environment does impact upon cognitive, perceptual or motor functions is not unequivocal. The lack of consistency in the quantification of ambient conditions, body core and skin temperatures restricts the value of many investigations. Differences in task duration and complexity may lead to equivocal conclusions. Overall, heat stress does appear to impact upon some forms of cognitive and motor performance.

Guidelines and procedures for selecting appropriate cognitive, perceptual and sustained attention tests are discussed. Four cognitive attributes are detailed by which the effects of heat on psychological performance in the field may be investigated. These are the attributes of vigilance, visual inattention, reasoning and time orientation. A further two attributes are suggested for laboratory studies: those of spatial orientation and auditory perception. Appropriate tests for determination of the effects of heat on psychological performance based upon the above attributes are recommended and described.

Experimental conditions detailing the degree of thermal strain appropriate for cognitive function tests in the heat are set out. A minimal strain of at least 38°C of body core temperature should be imposed. This strain should be achieved by a combination of the thermal environment and exercise and should be held for at least one hour prior to administration of the cognitive test.

Authors

Mark J. Patterson
University of Wollongong

Mark Patterson is a doctoral candidate in Environmental Physiology, Department of Biomedical Science, University of Wollongong. Past and present research projects include investigating the influence of upper-body skin temperatures on the control of sweat secretion, and the regulation of human body fluids during heat adaptation.

Nigel A.S. Taylor
University of Wollongong

Nigel Taylor is a Senior Lecturer in the Department of Biomedical Science, University of Wollongong specialising in Exercise and Environmental Physiology. He has collaborative research involvements with Aeronautical and Maritime Research Laboratory and the Naval Medical Research Institute (Bethesda, U.S.A.) which include human heat tolerance, cold-water adaptation and the provision of physiological specifications for breathing apparatus. Other research projects include investigations into human sweat regulation, heat adaptation, and the interaction of physical activity and ageing.

Denys Amos
Combatant Protection and Nutrition Branch

Denys Amos graduated from the University of Durham (UK) in 1960 with a BSc(Hons) and MSc (1961) in organic chemistry. He has worked with ICI and the Science Research Council and has been attached to CBDE in the UK. At AMRL he has undertaken extensive research into the decontamination and into protection of personnel against toxic chemicals. At present he is a Principal Research Scientist and manager of a program on personnel protection and physiological performance. He is also a Professorial fellow at the University of Wollongong. Recently he has been the principal investigator into the physiological assessments of soldier performance in the tropics and the newly developed Chemical, Biological Combat Suit.

Contents

1. Introduction

During exposure to hot environments, thermal homeostasis is transiently disrupted, with a resultant increase in stored body heat. To combat heat storage, mechanisms of heat dissipation are recruited; elevated skin blood flow, and secretion of eccrine sweat. If the rate of heat gain exceeds heat loss, the body will continue to store heat, resulting in an elevation in body core temperature (T_{core}).

Both physical and mental performance have been shown to be impaired in hot environments (Epstein, *et al.*, 1980; Nunneley *et al.*, 1982; Patterson *et al.*, 1994). The degree of impairment could possibly be related to a number of factors such as ambient temperature, relative humidity, radiant temperature, air movement, complexity of the task, acclimation/acclimatisation state, skill level of the subject, motivation, length of task, hydration level, and T_{core} and skin temperatures (T_{skin}). Hancock (1981) has suggested that cognitive performance is only impeded near the point of thermoregulatory collapse. However, most investigators argue that this interpretation may be misleading, and grossly underestimates the affect of heat on performance. Before providing recommendations concerning the selection and use of cognitive, perceptual and sustained attention (vigilance) tests to evaluate the effects of an elevation in T_{core} upon psychophysical function, we will briefly review a selection of the available literature on this topic. Before undertaking this, it is important to note that this report deals primarily with simple, single-function tests.

1.1 Literature Review

1.1.1 Vigilance

Mackworth (1950) was one of the first investigators to address the issue of thermal stress and human performance. A sustained attention task, involving the monitoring of a clock face for irregular double jumps in the revolving hand, was used to evaluate thermal influences. The task was performed in four different environments, each at a different effective temperature (21°C, 26°C, 31°C and 36°C). Performance deteriorated above at 31°C and 36°C for both response omissions and response time. In the other two temperatures, T_{core} was elevated, and was associated with decrements in performance. It was concluded that if the environment does not sufficiently alter T_{core}, then sustained attention performance may be maintained.

Benor and Shvartz (1971) assessed attention using an auditory vigilance task. Subjects walked on a treadmill at 3.5 km·hr^{-1} in an ambient temperature of 50°C. The relationship between missed signals and mean body temperature[1] displayed an exponential function, with a deflection point at approximately 38.5°C. The percentage of missed signals per minute of exposure rose steadily from 4% at a mean body

1 Mean body temperature is derived by the weighted combination of T_{core} and mean skin temperature (T_{skin}).

temperature of 37.5°C, to 12% at 38.5°C, from which missed signals rose sharply to 46%, at a mean body temperature of 39.0°C.

Some investigators have suggested that vigilance is not impaired in hot environments, with vigilance possibly even being improved (Loeb and Jeantheau, 1958; Fine *et al.*, 1960; Edholm, 1963; Colquhoun, 1969). For example, Colquhoun (1969) observed no change in signal omissions or response time between 27.8°C and 33.3°C effective temperature; although T_{core} was only elevated by 0.3°C. Similarly, Wilkinson *et al.* (1964) and Fox *et al.* (1963) found that auditory vigilance was optimal at a static T_{core} of 38.5°C.

1.1.2 Reaction Time

While vigilance has been shown to be diminished at an elevated T_{core}, response reaction time has been widely shown to be reduced (Epstein *et al.*, 1980; Nunneley *et al.*, 1982; Goodman *et al.*, 1984). However, differences in the effects of heat on reaction time have been found between simple reaction-time tests and serial-reaction tasks. The latter tasks generally involve a sustained attention component, where subjects respond once they observe random changes in serial stimuli. Reaction time is then recorded from the point of stimulus detection. Simple reaction-time tasks involve subjects performing responses after being alerted that a response is required.

Nunneley *et al.* (1982) used a simple reaction-time task, and observed that heat shortens reaction time. Goodman *et al.* (1984) made similar observations. They attributed this reduction in simple reaction time primarily to an elevation in intramuscular temperature, which acts to reduce motor-reaction time, while not affecting the pre-motor reaction time.

Fraser and Jackson (1955) observed an elevation in serial reaction-time to a visual vigilance task in a hot environment. Epstein *et al.* (1980) observed an optimal effective temperature of 30°C for producing the quickest reaction, while an effective temperature of 35°C increased the reaction time to the level experienced at an effective temperature of 21°C. However, Colquhoun (1969) observed no changes in reaction time to a vigilance task between effective temperatures of 27.8°C and 33.3°C. These unequivocal observations may indicate that the outcome of these experiments may have been influenced by the nature of the reaction-time task employed. Thus, simple reaction time appears to be reduced, while serial reaction time may remain unchanged or elevated.

1.1.3 Other Performance Tasks

Pursuit-rotor tasks have been used to determine the effect of heat load on co-ordination and motor performance. Allan and Gibson (1979) observed a decrement in performance at elevated levels of T_{core} and T_{skin}. Teichner and Wehrkamp (1954) found pursuit-rotor performance to be optimal at an environmental temperature of 21.1°C,

with an environmental temperature of 37.8°C reducing performance by over 30%. However, Sharma *et al.* (1986) found no alteration in coordination in a hot dry and hot humid environment in comparison to a thermoneutral environment. Their task involved subjects moving a stylus in a narrow groove, 0.5 cm wide, cut in an eight-cornered star shape. Epstein *et al.* (1980) showed a deterioration in aiming performance in a hot environment. Subjects were required to shoot at targets of different sizes. The greatest deterioration in aiming accuracy was recorded in an environment of 50°C and 40% relative humidity.

Both short and long-term memory have been shown not to be affected by an elevation in T_{core} and T_{skin} (Hollard *et al.*, 1985). Subjects were asked to recall a passage that they had learned 1 hour before heat exposure (long-term memory). Short-term memory was assessed using a digit span test, recalling a number series, both forwards and backwards. Similarly a reasoning task, involving a statement describing a letter sequence, with subjects responding 'true' or 'false', was not affected by environmental temperature. Two-digit subtractions were similarly unaffected by environmental conditions, however, the time in which the subtractions were completed was reduced at the elevated body temperatures. Sharma *et al.* (1986), however, found a reduction in performance on a running memory task in the hot environments, with the decrement being magnified in humid heat. The running memory task consisted of subjects being read a long list of numbers, at some point in time the number reading ceased and subjects were required to recall the last 5 numbers in reverse order. A substitution task, where subjects were required to match letters to geometric shapes, was not affected by the environmental conditions.

Bunnell and Horvath (1988) have observed similar results concerning short-term memory and reasoning. They used the Sternberg task, which presents 1-4 digits for 1 second followed by a single digit 1 second later. Subjects had to determine whether the single digit was a sub-set of the previous number set. This task was not affected by the environment, up to a wet bulb globe temperature of 30°C. Similarly, a divided-attention task (arithmetic), visual searching, and a tracking task were all unaffected by the hot environment.

Nunneley *et al.* (1982) used an orientation task to assess the effects of elevated body temperatures. A picture of a human form (manikin) held a circle in one hand and a square in the other, with either a circle or a square appearing at the bottom of the screen. Subjects had to determine which hand held the shape that appeared at the bottom of the screen. The manikin could be facing out of, or into the screen, and could be standing upright or upside down. Orientation was most affected at the highest elevation in body temperature. Conversely, Bunnell and Horvath (1988) found no change in this manikin task at elevated environmental conditions of 35°C (60% relative humidity), and 41°C (30% relative humidity). This may have been due to a failure to adequately elevate the T_{core} (see: Section 1.1.4).

Time orientation has also been evaluated. Such tasks require subjects to estimate designated time periods. Performance on such tasks has been shown to be impaired at high environmental temperatures (Bell, 1965; Fox *et al.*, 1967). Fox *et al.* (1967) observed that as T_{core} is elevated, the length of time estimated for a ten second period is reduced. That is, time orientation is affected, and subjects perceive the ten seconds to be a shorter period of time when their T_{core} is elevated.

Lockhart (1971) examined the effect of environmental temperature upon flicker-fusion threshold. Flicker fusion occurs when subjects can no longer detect the pulsatile nature of an oscillating light source, as the light starts to flicker at progressively higher frequencies. It was found that at elevated environmental temperatures, the flicker-fusion threshold was reduced, that is, the frequency of oscillation at fusion detection was higher. The investigator suggested that at higher ambient temperatures, neural excitability was elevated, therefore, increasing awareness of small visual alterations.

1.1.4 Confounding Factors

Body Temperature: A considerable amount of the human-factors research in this area has been undertaken by either physiologists, with a limited appreciation of the assessment of cognitive performance, or by psychologists, with a limited understanding of the physiological impact of the thermal environment. As a consequence, the literature is diluted by numerous poorly controlled experiments, and is therefore difficult to interpret. One of the major limitations of such research has been a failure to adequately quantify the thermal environment and the resultant changes in T_{core} and T_{skin}. Measuring only environmental conditions reduces the value of the results, since the thermal state of each subject is unknown.

Arees (1963) has postulated that cognitive performance may be related to the thermal gradients: core-skin and skin-environment. It was suggested that performance is better when the core-skin gradient is equivalent to that of skin-environment. That is, the outward flow of heat is unimpeded. However, when the core-skin gradient is small, heat is inadequately dissipated, and performance is hindered. In thermo-neutral conditions (air temperature = 25°C) T_{skin} is about 32°C, providing a skin-environment gradient of 7°C and a core-skin gradient of 5°C. In the heat, (air temperature = 38°C) the skin-environment gradient is reversed and reduced. Such thermal gradients may simply exert an affect via changes in heat storage, as seen through changes in T_{core} and T_{skin}, rather than as thermal gradient effects themselves.

Holland *et al.* (1985) observed no change in memory (short and long term), reasoning or mood state at an elevated T_{core} which was first elevated to 39.0°C, before testing began. However, there was a continual decrease in T_{core} during the performance of these cognitive tasks. The actual state of T_{core} with respect to performance, whether rising, falling or static, has received some attention. It has been suggested that stable, but elevated body temperatures may not induce significant differences in cognitive performance in hot environments (Hancock, 1986). Thermoreceptors sense

temperature changes, as well as static temperature, and are located at the skin and centrally within the body core. The signals emitted by these receptors are integrated at the hypothalamus, where an appropriate heat loss or gain mechanism is evoked. Afferent signals from these receptors are also conveyed to the cortex where sensations of warmth and cold are perceived. As temperature is changing, there is a large burst of activity. However, as the rate of local temperature change is reduced, such as that encountered when the new temperature level is approached, thermoreceptor firing rate is similarly reduced, with the frequency of neural impulses being considerably lower than produced by a given change in local temperature. Therefore, it may be quite understandable for investigators to observe different results at similar levels of T_{core}, as in one experiment the local temperature may be constant, whereas in the another it may maybe changing. This, of course, presupposes an inter-relationship between thermoreceptor function and psychophysical attributes. While such a relationship has been shown to exist for some attributes, such as thermal sensation, mood, perceived exertion and affective state, it has not been shown to exist for cognitive, perceptual or vigilance functions.

Allan and Gibson (1979) have perhaps used the most sophisticated design to address the difficulty of quantifying and controlling the thermal environment. They used a water-perfused garment to clamp T_{core} at three different levels (37.9°, 38.2° and 38.5°C), and performed a pursuit-rotor task at each level. The perfusion suit was first heated, so that T_{skin} was 38-39°C at each T_{core} level, then cooled so that T_{skin} was 35-36°C at each level of T_{core}. Performance was reduced at each level of T_{core} when the T_{skin} was elevated. It seemed that performance and thermal comfort tracked the changes in T_{skin}, and to a lesser degree T_{core}. It is possible that a reduction in T_{skin}, at an elevated T_{core}, warm receptor firing was diminished, resulting in greater thermal comfort at the same T_{core}.

Gibson *et al.* (1980) have attempted to determine if the rate of change in T_{core} and T_{skin} affects pursuit-rotor task performance. No statistical differences were observed, however, there did seem to be a trend within the data. The greater the rate of change in either T_{core} or T_{skin}, the lower the thermal comfort, and lower the ability of their subjects to perform the pursuit task.

The importance of T_{skin} in cognitive performance has been addressed by Nunneley *et al.* (1982) who focussed upon head T_{skin}. Cognitive performance was most impaired when both the head and body were being heated. Performance was actually enhanced when the head was cooled. It is difficult to conclude that head T_{skin} singularly improved performance, since T_{core} was elevated at a slower rate when the head was being cooled. Most probably, the improvement in performance resulted from the combined effects of a lower head T_{skin} and a more slowly changing T_{core}.

Task Complexity: The complexity of the cognitive, perceptual and attention tasks may account for the apparent thermal affects observed by different investigators. For example, Carlson (1961) varied task complexity in hot and neutral environments,

finding that performance was only reduced in the hot environment with the high complexity task. Similarly, Epstein *et al.* (1980) used three different sized aiming targets in three different environments (cool, moderately warm and hot). Aiming accuracy using the largest target was not affected by environmental temperature, however, as the target was reduced and environmental temperature was elevated, performance was compromised. The authors concluded that a difficult task reduced performance by 13%, heat stress reduced performance by 7%, and when the two stressors were combined, a magnified degradation of 27% was observed.

In a similar manner, the duration of the task may also account for differences between investigations (Wilkinson, 1969). This effect will be more pronounced in sustained attention (vigilance) tasks. For instance, Mortagy (1971) observed no reduction in performance in a hot environment when task duration was 20 min. However, when task duration was extended to 40 or 60 min, performance was impaired in a hot, compared with a neutral environment.

Hydration Status: Hydration state may also be a covariate, such that, when combined with thermal stress, it will magnify the effect of the environment upon performance. Evidence for this comes from work by Sharma *et al.* (1986). This group dehydrated subjects by 1, 2 and 3% of their body mass, and then examined cognitive and motor functions in thermoneutral and hot environments. Three tasks were completed: a substitution test, a running memory test and a coordination task (see: Section 1.1.3). While substitution was not significantly affected by either hydration or environment, there was a tendency for performance to decrease with increasing dehydration and environmental temperature. There was a critical level of 2% dehydration which impaired both running memory and coordination, with the reduction in performance being magnified in the hot environment. Gopinathan *et al.* (1988) has similarly highlighted a dehydration level of 2% at which psychological performance is diminished.

1.2 The effects of wearing clothing:

The above brief review does not include reference to the impact of clothing upon cognitive, perceptual or motor functions. Clearly, under some circumstances, the use of some clothing ensembles, such as nuclear, biological and chemical protective clothing, will have a strong impact upon performance within these domains. This topic has recently been extensively reviewed by Taylor and Orlansky (1993). The following points are worthy of note.

* Chemical warfare clothing will reduce manual dexterity, impede vision, degrade communication, increase respiratory stress, elevate psychological stress, reduce endurance time and the ability to work, and induce dehydration.
* Target detection, engagement times, and firing accuracy are degraded when wearing chemical warfare clothing.

* Chemical warfare clothing tends to be associated with changes in the psychological state of the wearer, with an intensification of symptom intensity and a general deterioration in mood state.

1.3 Conclusions

While there is some clear and convincing evidence that the thermal environment does have an impact upon cognitive, perceptual or motor functions, this evidence is not unequivocal. The lack of consistency in the quantification of ambient conditions, T_{core} and T_{skin} in many investigations restricts the value of such work. Similarly, differences in both task duration and complexity may lead to equivocal conclusions being drawn. To examine the effects of heat stress on cognitive performance, both T_{core} and T_{skin} need to be adequately measured and controlled. However, isolating the differential influence of T_{core} versus T_{skin} on cognitive performance would be a difficult task, since during most experimental designs, both tend to be altered in the same direction.

On the basis of the evidence reviewed above, it would appear that heat stress does impact upon some forms of cognitive and motor performance. These influences are apparent within: **sustained attention (vigilance); reaction time; spatial and time orientation.**

2. GUIDELINES AND PROCEDURES FOR SELECTING TESTS

The following section contains general guidelines and recommendations concerning the selection of cognitive, perceptual and sustained attention tests. It is the purpose of this section to outline considerations which are deemed of fundamental importance to any investigation designed to test the general hypothesis that heat strain impairs performance within these psychological domains.

2.1 The measurement design

The recommended test design for applied research in this field is the repeated measures model. Pre- and post-manipulation data are thus used to generate discrepancy scores, which are compared using standard statistical procedures. By chance alone, a certain amount of variation (scatter) can be expected to exist between the two test scores. However, in normal healthy subjects, this chance variation is random and relatively small (Cronbach, 1970). Thus, if significant discrepancies exist between sequential tests, then a pattern of functional deficit emerges. It is

recommended that this procedure, rather than normative comparison standards (compared with age-group norms) be used, since this provides a means of direct basal measurement for the variables of interest.

2.2 The purpose of the examination

In order to know the kind of information that should be obtained in a given series of tests, it is important to have a clear understanding of the purpose of the testing. That is, the assessor will select tests according to the nature of the group being assessed (the operational duties of the group), and the skills or characteristics deemed most appropriate to that group. Thus, the purpose of the examination may be determined by the group commander(s) themselves, or in consultation with scientific advisers. It is often necessary for the experimenter to then evaluate, and even to interpret these purposes before generating a test battery suited to an individual or group.

2.3 Formulation of experimental hypotheses

The result of this evaluation and interpretation should allow for the development of experimental hypotheses, which, in turn, lead to the selection of suitable functional tests. An overview of the operational hypotheses and test battery should then be returned to the originators of the requested information for verification, before assembly of the test battery commences.

2.4 Test selection considerations

Test selection may be driven by either the nature of the generated hypotheses (theory testing: Elmes *et al.*, 1989) or the suitability of the selected test(s) to the applied environment. Kantowitz (1992) suggests that both these criteria should be applied to the test selection process.

Let us consider the more pragmatic issue first. While one may choose a series of tests to evaluate one or more working hypotheses, it is possible that some of the tests chosen are unsuitable to the applied environment. This can occur when the scientific adviser has only a limited appreciation of the actual requirements of the research question, or the operational duties of the experimental group. For instance, one may select a test of vigilance which has very little relevance to the daily duties of the experimental group. In this instance, the test results, while permitting evaluation of generalised hypotheses related to vigilance, may not permit the derivation of group- or mission-specific outcomes. In this situation, it is essential that pilot assessment of selected tests be completed prior to commencing any experimental series. While exposure of the scientific adviser to the working environment can minimise the probability of inappropriate test selection or design, it cannot replace the need to trial tests in consultation with both the originators of the research question and the experimental group commander(s). In fact, in the applied field, one will often find feedback from the experimental group themselves to be a valuable component of the test selection

process. The limitation of using real tasks is that the results may become very difficult to interpret, and even lead to spurious conclusions (Lane *et al.*, 1986). Furthermore, there will be a limited ability of the researchers to generalise their observations to other situations, either applied or basic in nature.

The other purpose of such research is theory testing and the generalisation of derived outcomes across a broad range of duties. This type of testing may be performed within both the applied and basic research environments. Here the former shall be addressed. In selecting tests to evaluate the hypothesis that heat strain impairs neuromotor function, two general steps are recommended (adapted from: Kantowitz, 1992). First, conduct a field survey, in consultation with both the workers and their supervisors, to determine the broad types of skills required by the task under consideration, and group these skills into their corresponding psychophysical domains (Table 1). Second, using these classifications, it may be possible to undertake field tests (or observations) to identify those job tasks which are more susceptible to heat strain. From these observations (or pilot trials), it will be apparent that some attributes may be more affected than others. For instance, if it is shown that, for the duty of interest, tracking tasks are never performed under thermal stress conditions, then the inclusion of such a test will be of little relevance to either the experimental subjects or the perceived outcomes of the research. After identifying the relevant attributes which should be tested, select tests which are both valid and reliable (see: Sections 2.4.1 and 2.4.2), and which allow for testing of the hypotheses that heat strain impairs perceptual, cognitive or motor performance. Tests should be chosen that are frequently used, accepted within the scientific community, and generally well understood. Data derived from such work will not only have direct application to the applied question which generated the need for such research, but it will also have general application across other work-related tasks, as well as to the more basic body of scientific knowledge.

Table 1. Hypothetical breakdown of the psychophysical domains and job tasks which may be relevant to weapons operators.

Domain	*Job Tasks*
Perceptual domain	Visual interpretation
	Auditory cue detection
	Tactile recognition of controls
Cognitive domain	Sustained attention to clues
	Decision making
	Response selection
Motor domain	Response activation
	Reaction time

2.4.1 Test validity

The selection of individual test items must be based upon the validity of those items to hypotheses being tested. Since experimental hypotheses will vary between experimental groups, then the test batteries implemented to examine those hypotheses will similarly vary. Validity is defined as the ability of a test to measure or quantify specific predetermined attributes. A vigilance test with high validity will measure vigilance, however, one with low validity may rely upon arithmetic processes to evaluate vigilance. In such a case, the test may be a better index of arithmetic ability than it is of vigilance. There are several types of validity which need to be considered prior to test selection. These may be grouped into three broad categories: (i) content validity; (ii) criterion-related validity; and (iii) construct validity (Safrit, 1986).

In general, the applied researcher does not determine the validity of individual tests, since this has generally been performed by the test designers, with results being published in the scientific lierature. However, the applied scientist must still address this issue, since a test, validated on one group of subjects may no longer be valid for the experimental group of subjects. Violation of validity renders data interpretation difficult. In some applied situations, experimenters may need to establish very specific tests. For example, while the quantification of tracking ability may be valuable in determining the effects of heat on air traffic controller, it may be more appropriate to evaluate this ability using a realistic, finely controlled task, rather than a more standard measure of tracking ability. This is an acceptable alternative; however, since the test now differs from a more standard and validated tracking task, the experimenter can no longer compare experimental data across the two tasks.

(i) Content validity
This is quantification of the degree to which test items, or a test battery represents a defined field of content (American Psychological Association, 1984). For example, let us take a test battery designed to evaluate the general attribute motor performance. This general attribute may be sub-divided into some basic sub-components, which will vary according to how one perceives the task. However, one may identify some of the following sub-components: neuromotor function, visual scanning, tactile sensation, tactile memory and spatial orientation. A test battery with high content validity will permit evaluation of most, if not all of the identified sub-components. While this is a somewhat subjective process, content validity may be assessed in the following manner (Safrit, 1986):

* Examine the publisher's validity statement and table of specifications.
* Undertake the test yourself.
* Evaluate the field of content to determine: whether all the field sub-components are important to testing the proposed hypotheses; whether important sub-components have been omitted; whether unrelated sub-components have been inappropriately included; whether some sub-components receive inappropriate emphasis.

Content validity may be evaluated for self-developed tests, though this can be an arduous task. In the applied setting, it may be preferable to develop test batteries or test tasks which more closely replicate the working environment. However, consider closely the comments raised in Section 2.4.

Many gross motor performance tasks are difficult to evaluate using the criteria identified under content validity. Instead of utilising a collection of test items, such tests may use only a single test item, which is then repeated. In this circumstance, one must apply a test of logical validity. This is the extent to which the test evaluates attributes necessary to the performance of the specific task, for example, the routine duties of the experimental group. Logical validity may be assessed in the same manner as for content validity (Safrit, 1986).

(ii) Criterion-related validity
Numerous tests have been developed to evaluate single cognitive, psychophysical or motor attributes. For example, one can test tracking skills using several different tests. To assess their validity, one must compare such tests with some criterion reference, similar to the use of a calibration standard against which to compare physical measurements made within the laboratory. Before selecting a test, which itself is not a criterion test, one must determine whether its criterion-related validity has been established. That is, are the same experimental results obtained from both the criterion and the non-criterion tests. This is determined objectively using standard statistical procedures (*e.g.* the validity correlation coefficient). For most non-criterion tests, criterion-related validity will already have been established, and such data are readily available within the literature. As a general precaution, any restrictions which apply to the criterion-related validity of a test should be identified. For example, an intellectual ability test validated for use in pre-school children may no longer retain its validity with its application to adults. Similarly, a test of motor function may be invalidated when performed under conditions which impose physical constraints upon the subject, such as those encountered during the use of restrictive clothing. Under such circumstances the test may lose its criterion-related validity, and, in the process, become a test of the limits of clothing upon motor function, rather than of motor function *per se*.

For purpose-specific tests developed within applied research laboratories, the issue of criterion-related validity is much more complex. This form of validity may not be critical and it may be more important to evaluate the ability of an experimental group to perform tasks which most closely resemble their routine duties. Under such circumstances, criterion-related validity may be disregarded. However, such disregard does have a consequence. Since researchers can only draw conclusions related to the purpose-specific tests which they have developed and administered. While such tests may have cognitive, vigilance, visual and auditory functions contained within them, the experimenters will not be able to validly tease such components out of the test. This means that, while a performance decrement may be recorded, such work will not

allow for the identification of mechanisms underlying performance changes. Consequently, the test will be of limited value for designing intervention strategies. Furthermore, the results will have limited application to other experimental groups for whom the task in non-specific. This latter consideration renders such data of limited value to the scientific community at large.

(iii) Construct validity
This refers to the ability of the test to measure or quantify an attribute which cannot be directly measured and is often difficult to derive. Examples of such attributes or constructs include anxiety, work ethic and honesty. While the construct itself may not be able to be measured, various indicators of the construct can be quantified. Construct validity is generally determined by comparing the test scores of different groups. For example, within an anxiety test conducted while driving in simulated race conditions, one would expect people whose work or recreational pursuits regularly expose them to such stress would score lower on the anxiety indices than would learner drivers. An anxiety test failing to reveal group differences would be unlikely to possess construct validity.

2.4.2 Test reliability

Tests can be valid but unreliable. Similarly, tests can be reliable but invalid. While validity refers to the ability of the test to quantify the attribute of interest, reliability refers to the ability of the test to provide reproducible results. Ideally, tests are chosen which are both valid and reliable. Reliability may be determined by three standard methods: *(i) test-retest:* where simple correlation is used to determine test reliability; *(ii) single test administration:* where reliability is derived by comparing within test sections (*e.g.* split-half reliability); and *(iii) individual test score precision:* where the standard error of the measurement (when n is large) is used as an index of reliability.

2.5 Other considerations

2.5.1 Test presentation order

The sequence of presentation of tests within a battery does not have appreciable effects upon test performance (Cassel, 1962; Quereshi, 1968). An exception to this trend was noted by Neuger *et al.* (1981), when they observed that test of manual speed may be adversely affected when administered later in the day. Thus, as a general guide, it may advisable to administer the more difficult tests early in the battery. To minimise the possibility that test sequence affects the endeavour, diligence or attentiveness of the subjects, it may be advisable to alternate easy and more difficult tests. In this way subject attentiveness and concentration may be maintained, assuming that the subjects are motivated to work at peak capacity.

2.5.2 Testing the limits

Some of the tests which may be utilised require subjects to comprehend the full nature of the test, and then to complete the series of tasks dictated by the test. Some such tests may produce poor results, however, this reflects as much on the comprehension of the test requirements as on the performance of the subject on the test. In clinical practice, this may be evaluated by taking the patient beyond the limits of the test just administered, and asking the patient to complete the task outside the framework of the test itself. For example, in an arithmetic task, failed test items may be completed using pencil and paper, thereby enabling the experimenter to evaluate arithmetic ability, as opposed to test instruction comprehension. This procedure is recommended whenever the experimenter suspects that impairment of some function, other than that which is being evaluated, is interfering with the test performance. In the applied setting, such an assessment is made during pilot testing, and test instructions (written and verbal) are altered accordingly. However, the experimenter needs to be aware that, even after undertaking these precautions, some individuals will require additional instruction or clarification for them to undertake the test, and for the results to be interpreted relative to the hypotheses being tested.

2.5.3 Practice effects

As a general rule, tests which have a large speed component, have only one correct solution and involve unaccustomed response modes are susceptible to practice effects (Dodrill & Troupin, 1975). The effect of learning over multiple trials should be alleviated by familiarising and learning periods before evaluation, so that subjects attain a plateau in their performance prior to the experiment (Nunneley *et al.*, 1982). All chosen tests therefore should be evaluated for the impact of practice effects. Where such effects are found, serial trials should be completed to allow the experimenter to determine the nature of the learning curve. This curve may be modelled mathematically, with the number of repeat trials required to minimise practice effects being determined by the time taken to reach a given percentage (*e.g.* 95%) of the post-practice plateau. Experimenters are also encouraged to consider subject fatigue. While more relevant to complex tasks (Section 3.6), than to short-duration, simple-function assessments, fatigue will have a strong impact upon performance in the field.

2.5.4 Experimental conditions

The experimental conditions which facilitate hypothesis testing must be carefully selected. First, basal values must be established which ensure that the performance of the experimental subject is optimised. These are typically standardised and reproducible conditions. Some tests have standardised conditions prescribed which cover environmental factors and test lighting, audio-visual distraction, test presentation style, provision of knowledge of results, instructions on word usage and supplementary explanations. The basal data collection must also allow for the

controlling of confounding factors. For example, if testing is to be performed with the subjects wearing protective clothing, then they should wear such clothing during the control conditions.

During the experimental trials, the subject sample and the environment should be optimised to best achieve the desired conditions (Kantowitz, 1992). While, in general, it is often valuable to use subjects drawn from the population to which the experimental results will be applied, this need not be maintained rigidly. If the purpose of the research is theory testing, then this guideline may be ignored. For example, if one is testing the hypothesis that heat strain impairs vigilance (see: Section 3.5), then it is sufficient to use a homogenous sample drawn from the greater population. Results obtained from such work should be able to be applied generally to a variety of applications. However, if the applied question simply relates to the role of heat strain on job-specific duties, then samples must be drawn from the population of people currently trained in those duties. The limitation of data derived from such testing is that it is not easily generalised either to other applications or to more theoretical frameworks (Kantowitz, 1992).

Optimisation of the environment is more complex. Consideration must be given to a variety of physical components which combine to make up the working environment (*e.g.* noise, lighting, vibration, thermal stress). Again, either theoretical or pragmatic outcomes can be used to determine the relative importance of these components within the experimental setting. When testing theory, the laboratory is modified to replicate and control the most important characteristics of the working environment. However, field testing is often best used when purely practical outcomes drive the research. For the purpose of this report, only the thermal environment will be addressed.

It is necessary to determine, *a priori,* whether it is of interest to evaluate the effects of thermal stress *per se,* or how such stress impacts upon thermal strain[2]. While the distinction may be obvious to those familiar with such experiments, it is apparent within the literature, that this has not been universally addressed. Thus some groups, in attempts to address this issue, have simply exposed subjects to heat stress, without quantifying the concomitant heat strain. Assuming that heat strain is the real concern, then it must be determined how the required thermal load may best be imparted to the subjects. This is a dual issue, since it involves the possible inclusion of exogenous (environmental) and endogenous (metabolic) thermal stresses. Such work will be driven, at least in the first instance, by applied motives, so it becomes important to determine how the heat stress may best be achieved. Since the nature of the thermal stress dictates, to a large extent, the nature of the physiological response, the heat stress may best be achieved by replicating realistic working environments. Accordingly, prescriptive details are required for dry, wet and black globe

---

2 Heat stress refers the physical properties of the environment (air temperature, relative humidity, radiant heat load), while heat strain quantifies the magnitude of the physiological responses to this stress.

temperatures, and wind velocity. Similarly, the nature of the thermal strain requires definition. This may be defined as a set T_{core}, a rate of T_{core} change, or some combination of T_{core} and T_{skin}. Since the duration spent under heat loading will affect physiological function and fatigue, this feature must be rigidly controlled, while also allowing for the attainment of a significant thermal strain. Considering these points the following recommendations are suggested:

(1) A minimal strain of **at least 38°C (T_{core})** be imposed.
(2) This strain be achieved by a **combination of the thermal environment and exercise.**
(3) This strain be held for **at least 1 hour** prior to test administration, to ensure the attainment of a steady thermal state and thermal equilibrium between tissue beds.

3. RECOMMENDED DOMAINS FOR TESTING

Fourteen cognitive, perceptual and attention tests have been outlined and recommended for use. Tests from a variety of domains are listed. This list is not exhaustive and there is a wide variety of such tests available within the literature. A detailed search of the literature is advisable before constructing test batteries. The inclusion of test domains, and tests within these domains, was based upon their perceived appropriateness to some requirements and activities of the Australian Defence Force. It is important to note that some of the recommended tests are drawn directly from research related to brain dysfunction and may be somewhat insensitive when applied to normal populations.

3.1 Perceptual-function tests

3.1.1 Visual-inattention

Visual-inattention phenomenon[3] relates to the absence of awareness of visual stimuli which occur in the left field of vision. Visual inattention is, therefore, associated with right hemisphere dysfunction. Since it is the right hemisphere which dominates in the processing of visual information, and generally dominates the attention domain, a loss of visual acuity in the left field of vision corresponds with reduced visual attention.

Line-bisection test: The multiple trial test version developed by Schenkenberg *et al.* (1980) is recommended. Subjects are presented with 20 lines of different lengths, some of which cross the midline of the page. Six of these lines are centred to the left of the midline, six to the centre, and six to the right of the midline. The top and bottom lines

[3] Also referred to as visual extinction or visual neglect.

are centred. The subject is asked to "cut each line in half by placing a small pencil line through each line, as close as possible to the centre". Other points: the non-drawing hand is kept off the table; only one mark is to be made per line; no lines are to be skipped; lines are marked in sequence; the trial is repeated with non-dominant hand with the page rotated 180° for this trial. Two scores are obtained: (i) the number and position of unmarked lines (*e.g.* 1R, 0C, 3L); (ii) the percent deviation score, which quantifies the extent to which the subject failed to correctly estimate the true centre of each line (percent deviation = [measured left half-true half]/true half*100): positive scores recorded when the right hand is used are indicative of the visual inattention phenomenon.

[Schenkenberg, T., Bradford, D.C., and Ajax, E.T. (1980). Line bisection and unilateral visual neglect in patients with neurologic impairment. *Neurology*. 30:509-517.]

3.1.2 Visual recognition

Visual recognition, which requires processing and storage of visual data, is also a right hemisphere attribute. Thus, poor performance in these recognition tests may be interpreted as indicating an interference or dysfunction in right hemisphere function.

Judgement of line orientation: Subjects are required to estimate angular relationships between line segments. Eleven numbered lines are arranged to form a semicircle. Paired lines are then presented to the subjects, who are required to determine which of the numbered lines are being represented. Thirty items are administered in a single test.

[Benton, A.L., Varney, N.R., and Hamsher, K. de S. (1978) Visuospatial judgment. A clinical test. *Archives of Neurology*. 35:364-367.]

3.1.3 Visual organisation

Subjects are required to make sense out of sectioned, incomplete, vague and distorted visual stimuli (Elmes *et al.*, 1989). Such a task requires some degree of perceptual recognition, but extends this requirement into perceptual organisation. Again, this attribute is dominated by the right hemisphere.

The Hooper visual-organisation test: The test consists of thirty pictures, representing cut-up and dispersed pieces of common objects. The nature of the pictures varies with changes in the degree of difficulty. Verbal or written responses may be given.

[Hooper, H.E. (1958). The Hooper Visual Organization Test. Manual. Los Angeles: Western Psychological Services.]

3.1.4 Visual scanning

Visual scanning defects will compromise basic skills such as reading, writing and telling time (Diller *et al.*, 1974). While it is most unlikely that thermal strain will affect scanning to any great extent, the relevance of this attribute to various military personnel warrants its inclusion.

Perceptual maze test: A lattice-type maze, triangular in shape, with randomly placed points, requires subjects to trace from the bottom to the top through as many points as possible in one minute. There are 18 different mazes, each having its own normative data set. In this task the subject not only has to use perceptual abilities but must also be able to comprehend a rather complex task, count, keep track of several numbers and the paths they represent and choose between alternate routes.

[Elithorn, A., Jones, D., Kerr, M., and Lee, D. (1964). The effects of the variation of two physical parameters on empirical difficulty in a perceptual maze test. *British Journal of Psychology*. 55:31-37.]

3.1.5 Non-verbal auditory perception

The dominant hemisphere for verbal function (reading, writing, speaking, verbal memory) is the left hemisphere. Thus, auditory tests are a useful means by which left hemisphere function can be differentiated. While there is no inherent reason to suspect functional differences between the hemispheres during thermal strain, it is recommended that test batteries include both visual and auditory tests. Since many military tasks involve the use of non-verbal auditory stimuli, these tests also appear to have a strong practical significance.

Seashore rhythm test: As the name suggests, this test requires subjects to discriminate between like and unlike pairs of musical rhythms. For example, a series of three evenly spaced taps is first presented, followed by a second series of three taps, with the first two taps being closer together, and a slight delay between the second and third taps. Subjects determine whether the rhythms are identical.

[Seashore, C.E., Lewis, D., and Saetveit, D.L. (1960). *Seashore measures of musical talents.* (Rev. ed.). New York: Psychological Corporation.]

3.2 Memory-function tests

3.2.1 Digit recall

Wechsler memory scale digit-span test: This short-term memory task involves subjects being read digits at a frequency of one per second. The number of digits presented can vary, to increase or decrease difficulty. Digits are then recalled when requested, with the order of recall being either in a forwards or backwards sequence.

[Wechsler, D. (1955). Wechsler Adult Intelligence Scale. Manual. New York: Psychological Corporation.]

3.2.2 Visual memory

Visual retention test: This is a multiple-choice recognition task. Twenty white cards are available for presentation to the subject, each containing four blackened squares variously positioned. The positioning of these squares is such that the shapes so formed, are different on each of the twenty white cards. A stimulus card is exposed for two seconds. The subject identifies the stimulus card from a set of four similar cards. A second stimulus lasting for ten seconds is provided, with the card rotated by 180°. Error scores are recorded.

[Warrington, E.K., and James, M. (1967). Disorders of visual perception in patients with localized cerebral lesions. *Neuropsychologia.* 5:253-266.]

3.3 Conceptual function tests

3.3.1 Verbal reasoning problems

Reasoning tests require various forms of logical thinking, an understanding of relationships between information, and some degree of practical judgement.

Poisoned-food problem: Ten problems, and a practice problem, are presented to subjects. Subjects receive a work sheet with nine foods listed, and columns for information relating to the meal and whether the person consuming the meal lived or died. The ten problems are given to each subjects. The task is to identify which food caused death in each of the problems provided.

[Arenberg, D. (1968). Concept problem solving in young and old adults. *Journal of Gerontology.* 23:279-282.]

3.3.2 Arithmetic problems

Arithmetic reasoning problems: Little mathematical skill is required, as subjects need to make comparisons between elements of the problem. An example would be, "The green basket contains three apples; the blue basket has twice as many. How many apples are there altogether?". This is quite simple, however the level of difficulty can be increased.

[Luria, A.R. (1973). The working brain: an introduction to neuropsychology. New York: Basic Books.]

3.4 Orientation tests

An appreciation of orientation[4] requires consistent and dependable integration of attention, perception and memory. This strong reliance upon various processes makes orientation exceedingly vulnerable to the effects of brain dysfunction (Schulman *et al.*, 1965). It is therefore suggested that such tests may be well suited to testing the effects of thermal strain.

3.4.1 Time orientation

Time estimation tests: Subjects estimate the passage of selected time intervals, ranging from 10 seconds to 5 minutes. The most commonly used time period is a one minute period. Another method of evaluating time estimation requires subjects estimating the length of time taken to complete a given test session.

[Benton, A.L., Van Allen, M.W., and Fogel, M.L. (1964). Temporal orientation in cerebral disease. *Journal of Nervous and Mental Disease.* 139:110-119.]

3.4.2 Space orientation

Mental re-orientation: Figures of men which hold disks in their hands are presented to the subjects, with one disk being black. The men have four different standing positions: facing forwards; facing backwards; standing upright; standing upside down. Each position is shown four times with black disks being equally distributed between the two hands. Subjects need to indicate which hand is holding the black disk.

[Ratcliff, G. (1979). Spatial thought, mental rotation and the right cerebral hemisphere. *Neuropsychologia.* 17:49-54.]

3.5 Attention tests

Attention deficits, in their purest form, are manifest as a reduced ability to focus on a given task. Such deficits may be induced by attention disturbances, which can be simply identified in sustained attention tasks (vigilance). However, decreased vigilance will impair performance on more complex tasks, such as those requiring conceptual tracking. While complex attention tasks are available, we recommend the more simple tests, since they permit an easier identification of the mechanisms leading to attention deficit. Some more complex sustained attention tasks may be adversely affected by changes in attributes other than sustained attention, thereby making data interpretation difficult. However, it is important to note that such tasks are often of greater practical significance.

---

4 The perception of oneself in relation to the surrounding environment, objects within that environment, and events occurring within that environment.

3.5.1 Sustained attention test (vigilance)

Letter cancellation test: Subjects are presented with a sheet of paper containing 16 rows of 26 single-spaced lower case letters, interspersed with ten capitals and four double spaces. Tests can be at different levels: (a) cross out capitals; (b) cross out capitals and letters following a double space; and (c) cross out capitals and letters preceding double space. Scoring can be for speed, errors and omissions. Task variations have been used, with subjects required to cross out specific letters rather than capitals.

[Talland, G.A., and Schwab, R.S. (1964). Performance with multiple sets in Parkinson's disease. *Neuropsychologia*. 2:45-53.]

3.5.2 Tracking tests

Paced auditory serial addition test: Sixty pairs of random digits are read to the subject. The subject is required to add each digit pair once he/she hears the stimulus digit. If the digits read were "5-4-6-8-9", and the stimulus digit was "4", then the correct response would be "9-10-14-17". The subject would begin as soon as the digit "4" sounded. The digits can be presented at different rates (1.2-2.4 seconds between digits), with performance being evaluated in terms of the percentage of correct responses or a mean score.

[Gronwall, D.M.A., and Sampson, H. (1974). *The psychological effects of concussion.* Auckland, N.Z.: Auckland University Press/Oxford University Press.]

3.5.3 Complex attention functions

Symbol digit modalities test: A simple substitution task. To each presented symbol, the subject uses a substitution number, which is taken from a corresponding legend of symbols and numbers. The response may be written or verbal. A set time period of 90 seconds is allowed for completion of 110 items.

[Smith, A. (1968). The Symbol Digit Modalities Test: a neuropsychologic test for economic screening of learning and other cerebral disorders. *Learning Disorders*. 3:83-91.]

3.6 Combined or complex tasks

A number of tests found within the literature have used both combined and complex tasks. For example, more than one attribute is tested within a single task. While these tasks are both novel and challenging, the underlying attribute causing the reduction in performance cannot be readily determined due to interaction or interference of other attributes included within the test. For example, Sharma *et al.* (1986) have described a task which they have termed a *concentration test.* The test consisted of long series of

numbers being read to subjects at a rate of one per second. At some point in time, the reading ceased, and the subject was required to read out the last five numbers in reverse order. This task tests both attention and running memory. Therefore, if performance was hindered, the resultant cause of the decrement would be unknown. Similarly, Thompson (1973) has employed a repeated-acquisition task, requiring the learning and recall of chains of movement (behavioural) sequences using a cursor on an illuminated display. We have chosen to not review these combined or complex tasks, since the results obtained from such tests give no indication of which process is impeded during a given experimental state.

Other investigators have deliberately included two tasks running simultaneously in their test battery (*e.g.* Provins and Bell, 1970). A serial reaction time task was combined with a visual vigilance task to assess the effects of heat on performance. The serial reaction time task consisted of five lights positioned directly in front of the subject. When a light came on, the subject responded by hitting one of five buttons located in front of the body. The pace of the serial reaction time task was varied (either slow or fast). The visual vigilance task, which ran at the same time as the serial reaction time, involved subjects turning off lights, with the use of six foot switches, once the corresponding light was illuminated. Six lights were positioned in a semi-circle, from 87° left of the subject to 87° right of the subject, at a distance of about 2 m. While these combined tasks add a new level of novelty and complexity to testing procedures, deterioration in performance may be difficult to interpret. For example, visual extension (dominated by left sided visual dysfunction) may interfere with the latter task, rendering its value as a vigilance task difficult to assess.

4. RECOMMENDED COGNITIVE, PERCEPTUAL AND ATTENTION TESTS

The following Table summarises our recommendations for tests from the cognitive, perceptual and attention domains which, in the opinion of the authors, are well suited for use in thermal stress experiments, and which meet the guidelines outlined with Section 2. Each test has been briefly described within Section 3, and we have recommended tests for both field- and laboratory-based research. The recommendations are based on satisfying the guidelines outlined in Section 2 and on perceived relevance to military applications. However, it should be remembered that this test list is neither exhaustive nor exclusive and readers are advised to consult both the available literature, AGARD (1989) and DPSYCH-A before test batteries are finalised.

Table 2. Recommended tests to determine the effects of heat on mental performance.

Attribute	*Task (Section No.)*	*Field testing*	*Laboratory Testing*
Vigilance	3.5.1	*	*
Visual Inattention	3.1.1	*	*
Reasoning	3.3.1	*	*
Time Orientation	3.4.1	*	*
Spatial Orientation	3.4.2		*
Auditory Perception	3.1.5		*

5. REFERENCES

AGARD. (1989). Human performance assessment methods. NATO Advisory Group for Aerospace Research and Development. Report No. AGARD-AG308.

Allan, J.R., and Gibson, T.M. (1979). Separation of the effects of raised skin and core temperature on performance of a pursuit rotor task. *Aviat. Space Environ. Med.* 50:678-682.

American Psychological Association. (1984). *Joint standards for educational and psychological tests.* American Psychological Association. Washington, D.C.

Arees, E.A. (1963). *The effect of environmental temperature and altering stimuli on prolonged search.* Unpublished doctoral dissertation, University of Massachusetts, Amherst.

Bell, C.R. (1965). Time estimation and increases in body temperature. *Journal of Experimental Psychology.* 70:232-234.

Benor, D., and Shvartz, E. (1971). Effect of body cooling on vigilance in hot environments. *Aerospace Medicine.* 42:727-730.

Bunnell, D.E., and Horvath, S.M. (1988). Interactive effects of heat, physical work, and CO exposure on metabolism and cognitive task performance. *Aviat. Space Environ. Med.* 60:428-432.

Carlson, L.D. (1961). Human performance under different thermal loads. (Tech. Rep. No. 61- 43). Brooks AFB, TX:U.S.A.F. Aerospace Medical Center, School of Aviation Medicine.

Cassel, R.H. (1962). The order of the tests in the battery. *Journal of Clinical Psychology.* 18:464-465.

Colquhoun, W.P. (1969). Effects of raised ambient temperature and event rate on vigilance performance. *Aerospace Medicine.* 40:413-417.

Cronbach, L.J. (1970). *Essentials of psychological testing.* N.Y.: Harper & Row.

Diller, L., Ben-Yishay, Y., Gerstman, L.J., Goodkin, R., Gordon, W., and Weinberg, J. (1974). *Studies in cognition and rehabilitation in hemiplegia* (Rehabilitation Monograph No. 50). New York: New York University Medical Centre Institute of Rehabilitation Medicine.

Dodrill, C.B., and Troupin, A.S. (1975). Effects of repeated administrations of a comprehensive neuropsychological battery among chronic epileptics. *Journal of Nervous and Mental Diseases.* 161:185-190.

Edholm, O.G. (1963). Heat acclimatization studied in the laboratory and the field. A multi-disciplinary approach. *Ergonomics.* 6:304-305.

Elmes, D.G., Kantowitz, B.H., and Roediger, H.L. (1989). *Research methods in psychology.* West Publ. St. Paul, MN.

Epstein, Y., Keren, G., Moisseiev, J., Gasko, O., and Yachin, S. (1980). Psychomotor deterioration during exposure to heat. *Aviat. Space Environ. Med.* 51:607-610.

Fine, B.J., Cohen, A., and Crist, B. (1960). Effect of exposure to high humidity at high and moderate ambient temperature on anagram solution and auditory discrimination. *Psychological Reports.* 7:171-181.

Fox, R.H., Goldsmith, R., Hampton, I.F.G., and Wilkinson, R.T. (1963). The effect of a raised body temperature on the performance of mental tasks. *Journal of Physiology*. 166:22P-23P.

Fox, R.H., Bradbury, P.A., and Hampton, I.F.G. (1967). Time judgment and body temperature. *Journal of Experimental Pyschology*. 75:88-86.

Fraser, D.C., and Jackson, K.F. (1955). Effect of heat stress on serial reaction time in man. *Nature*. 176:976-977.

Gibson, T.M., Redman, P.J., and Allan, J.R. (1980). Effect of direction and rate of change of deep body and skin temperatures on performance of a rotary pursuit task. *Aviat. Space Environ. Med.* 51:445-447.

Goodman, D., Hancock, P.A., Runnings, D.W., and Brown, S.L. (1984). Temperature-induced changes in neuromuscular function: central and peripheral mechanisms. *Perceptual and Motor Skills*. 59:647-656.

Gopinathan, P.M., Pichan, G., and Sharma, V.M. (1988). Role of dehydration in heat stress-induced variations in mental performance. *Arch. Env. Health*. 43:15-17.

Hancock, P.A. (1981). Heat stress impairment of mental performance: a revision of tolerance limits. *Aviat. Space Environ. Med.* 52:177-180.

Hancock, P.A. (1986). Sustained attention under thermal stress. *Psychological Bulletin*. 99:263- 281.

Hohnsbein, J., Piekarski, C., Kampmann, B., and Noack, Th. (1984). Effects of heat on visual acuity. *Ergonomics*. 27:1239-1246.

Holland, R.L., Sayers, J.A., Keatinge, W.R., Davis, H.M., and Peswani, R. (1985). Effects of raised body temperature on reasoning, memory, and mood. *J. Appl. Physiol*. 59:1823-1827.

Kantowitz, B.H. (1992). Selecting measures for human factors research. *Human Factors*. 34:387-398.

Lane, N.E., Kennedy, R.S., and Jones, M.B. (1986). Overcoming the unreliability of operational measures: the use of surrogate measure systems. In: *Proceedings of the Human Factors Society 30th annual meeting*. Santa Monica, CA. Pp. 1398-1402.

Leob, M., and Jeantheau, G. (1958). The influence of noxious environment stimuli on vigilance. *Journal of Applied Psychology*. 42:47-49.

Lockhart, J.M. (1971). Ambient temperature and the flicker-fusion threshold. *Journal of Experimental Pyschology*. 87:314-319.

Mackworth, N.H. (1950). Researches on the measurement of human performance. *Medical Research Council Special Report Series 268*. London: Her Majesty's Stationary Office.

Mortagy, A.K. (1971). *Effects of work/rest schedules on monitoring performance in the heat*. Unpublished doctoral dissertation, Texas Tech University, Lubbock.

Neuger, G.J., O'Leary, D.S., Fishburne, F., et al. (1981). Order effects on the Halstead-Reitan Neuropsychological Test Battery and allied procedures. *Journal of Consulting and Clinical Psychology*. 49:722-730.

Nunneley, S.A., Reader, D.C., and Maldonado, R.J. (1982). Head-temperature effects on physiology, comfort, and performance during hyperthermia. *Aviat. Space Environ. Med.* 53:623-628.

Patterson, M.J., Warlters, D., and Taylor, N.A.S. (1994). Attenuation of the cutaneous blood flow response during combined exercise and heat stress. *Eur. J. Appl. Physiol.* 69:367- 369.

Provins, K.A., and Bell, C.R. (1970). Effects of heat stress on the performance of two tasks running concurrently. *Journal of Experimental Psychology*. 85:40-44.

Quereshi, M.Y. (1968). The comparability of WAIS and WISC subtest scores and IQ estimates. *Journal of Clinical Psychology*. 68:73-82.

Safrit, M.J. (1986). *Introduction to measurement in physical eduction and exercise science.* Mosby College Publ. St Louis.

Schulman, J.C., Kaspar, J.C., and Throne, F.M. (1965). *Brain damage and behavior.* Springfield, Ill: C.C. Thomas.

Sharma, V.M., Sridharan, K., Pichan, G., and Panwar, M.R. (1986). Influence of heat-stress induced dehydration on mental functions. *Ergonomics*. 29:791-799.

Taylor, H.L., and Orlansky, J. (1993). The effects of wearing protective chemical warfare combat clothing on human performance. *Aviat. Space & Environ. Med.* 64:A1-A41.

Teichner, W.H., and Wehrkamp, R.F. (1954). Visual-motor performance as a function of short-duration ambient temperature. *Journal of Experimental Psychology*. 47:447-450.

Thompson, M.D. (1973). Repeated acquisition as a behavioural baseline for studying drug effects. J. Pharmacol. Exp. Ther. 184: 506-514.

Wilkinson, R.T., Fox, R.H., Goldsmith, R., Hampton, I.F.G., and Lewis, H.E. (1964). Psychological and physiological responses to raised body temperature. *J. Appl. Physiol.* 19:287-291.

Wilkinson, R.T. (1969). Some factors influencing the effect of environmental stressors upon performance. *Psychological Bulletin*. 72:260-272.

DISTRIBUTION LIST

Tests of cognitive, perceptual and sustained attention functions in hot environments.

M.J. Patterson, N.A. S. Taylor and D. Amos

AUSTRALIA

DEFENCE ORGANISATION

Task Sponsor
DGDFHS

S&T Program
Chief Defence Scientist
FAS Science Policy
AS Science Corporate Management
} shared copy
Director General Science Policy Development
Counsellor Defence Science, London (Doc Data Sheet)
Counsellor Defence Science, Washington (Doc Data Sheet)
Scientific Adviser to MRDC Thailand (Doc Data Sheet)
Director General Scientific Advisers and Trials/Scientific Adviser Policy and Command (shared copy)
Navy Scientific Adviser (Doc Data Sheet and distribution list only)
Scientific Adviser - Army
Air Force Scientific Adviser
Director Trials

ESRL
Chief of Land Operations Division
Dr J. Manton
Mr J. Hansen

Aeronautical and Maritime Research Laboratory
Director
Head CPNB
Dr S. Oldfield
Author(s):
Mr M.J. Patterson (University of Wollongong)
Dr N.A.S. Taylor (University of Wollongong)
Prof. D. Amos
Dr W.M. Lau

DSTO Library
Library Fishermens Bend
Library Maribyrnong
Library Salisbury (2 copies)
Australian Archives
Library, MOD, Pyrmont (Doc Data sheet only)

Capability Development Division

Director General Maritime Development (Doc Data Sheet only)
Director General Land Development
Director General C3I Development (Doc Data Sheet only)

Navy

SO (Science), Director of Naval Warfare, Maritime Headquarters Annex, Garden Island, NSW 2000 (Doc Data Sheet only)

Army

DGLD
ABCA Office, G-1-34, Russell Offices, Canberra (4 copies)
SO (Science), DJFHQ(L), MILPO Enoggera, Queensland 4051
NAPOC QWG Engineer NBCD c/- DENGRS-A, HQ Engineer Centre LiverpoolMilitary Area, NSW 2174

Intelligence Program

DGSTA Defence Intelligence Organisation

Corporate Support Program (libraries)

OIC TRS, Defence Regional Library, Canberra
Officer in Charge, Document Exchange Centre (DEC), 1 copy
*US Defence Technical Information Center, 2 copies
*UK Defence Research Information Centre, 2 copies
*Canada Defence Scientific Information Service, 1 copy
*NZ Defence Information Centre, 1 copy
National Library of Australia, 1 copy

UNIVERSITIES AND COLLEGES

Australian Defence Force Academy
Library
Head of Aerospace and Mechanical Engineering
Deakin University, Serials Section (M list), Deakin University Library, Geelong, 3217
Acquisition Librarian, University of Wollongong
Senior Librarian, Hargrave Library, Monash University
Librarian, Flinders University
James Cook University (Dr R. Griffiths)

OTHER ORGANISATIONS

NASA (Canberra)
AGPS
NT Institute of Sports (Dr D. Hatcher)

OUTSIDE AUSTRALIA

ABSTRACTING AND INFORMATION ORGANISATIONS

INSPEC: Acquisitions Section Institution of Electrical Engineers
Library, Chemical Abstracts Reference Service
Engineering Societies Library, US

Materials Information, Cambridge Scientific Abstracts, US
Documents Librarian, The Center for Research Libraries, US

INFORMATION EXCHANGE AGREEMENT PARTNERS

Acquisitions Unit, Science Reference and Information Service, UK
Library - Exchange Desk, National Institute of Standards and Technology, US

SPARES (10 copies)

Total number of copies: 64

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION
DOCUMENT CONTROL DATA

1. PRIVACY MARKING/CAVEAT (OF DOCUMENT)

2. TITLE	3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION)
Tests of cognitive, perceptual and sustained attention functions in hot environments	Document (U) Title (U) Abstract (U)
4. AUTHOR(S) Mark J. Patterson, Nigel A.S. Taylor and Denys Amos	5. CORPORATE AUTHOR Aeronautical and Maritime Research Laboratory PO Box 4331 Melbourne Vic 3001 Australia

6a. DSTO NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DSTO-TR-0650	AR-010-503	Technical Report	September 1997

8. FILE NUMBER	9. TASK NUMBER	10. TASK SPONSOR	11. NO. OF PAGES	12. NO. OF REFERENCES
510/207/0847	ADF 96/065	DGDFHS	25	44

13. DOWNGRADING/DELIMITING INSTRUCTIONS	14. RELEASE AUTHORITY
	Head, Combatant Protection and Nutrition Branch

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved for public release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE CENTRE, DIS NETWORK OFFICE, DEPT OF DEFENCE, CAMPBELL PARK OFFICES, CANBERRA ACT 2600

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CASUAL ANNOUNCEMENT Yes

18. DEFTEST DESCRIPTORS

Performance tests, Cognition, Cognitive ability, Motor ability, Perception, Performance (human), Hot weather, Heat stress.

19. ABSTRACT

The report reviews the utility of various tests of cognitive function during human performance in hot conditions. The evidence that the thermal environment does impact upon cognitive, perceptual or motor functions is not unequivocal. The lack of consistency in the quantification of ambient conditions, body core and skin temperatures restricts the value of many investigations. Differences in task duration and complexity may lead to disputable conclusions being drawn. Overall, heat stress does appear to impact upon some forms of cognitive and motor performance. Guidelines and procedures for selecting appropriate cognitive, perceptual and sustained attention tests are discussed. Tests suitable for determination of the effects of heat on psychological performance are recommended. Experimental conditions detailing the degree of thermal strain appropriate for cognitive function tests in the heat are described.